Final Report for Period: 04/2010 - 03/2011
Submitted on: 05/17/2011
Principal Investigator: Dellaert, Frank .
Award ID: 0448111
Organization: Georgia Tech Research Corp
Submitted By:
Dellaert, Frank - Principal Investigator
Title:
CAREER: Markov Chain Monte Carlo Methods for Large Scale Correspondence Problems in Computer Vision and Robotics

Project Participants

Senior Personnel

Name: Dellaert, Frank
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Khan, Zia
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Potthast, Christian
Worked for more than 160 Hours: Yes
Contribution to Project:

Post-doc

Name: Mariottini, Gian-Luca
Worked for more than 160 Hours: Yes
Contribution to Project:

Graduate Student

Name: Kaess, Michael
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: McEwan, Hunter
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Quennesson, Kevin
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Ranganathan, Ananth
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Schindler, Grant
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Kamal, Shakti
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Jones, Josh
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Rogers, John
Worked for more than 160 Hours: Yes
Contribution to Project:
Incoming PhD student, working on 3D-2D correspondence

Name: Krauthausen, Peter
Worked for more than 160 Hours: Yes
Contribution to Project:
Unfunded M.Sc. student doing exploratory research with the PI

Name: Kipp, Alexander
Worked for more than 160 Hours: Yes
Contribution to Project:
Unfunded M.Sc. student doing exploratory research with the PI

Name: Oh, Sang Min
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Ni, Kai
Worked for more than 160 Hours: Yes
Contribution to Project:
Supported from project in Fall 2007 as Graduate Research Assiatant

Name: Sun, Mingxuan
Worked for more than 160 Hours: Yes
Contribution to Project:

Undergraduate Student

Name: Cheng, Tom
Worked for more than 160 Hours: Yes
Contribution to Project:

Name: Berger, Justin
Worked for more than 160 Hours: Yes
Contribution to Project:

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

DSTA, Singapore

? DSTA, Singapore, has provided a related seed-grant at a rate of $30K per year, over 3 years

Microsoft Research

? Microsoft Research has supported the research with a gift of $20K

University of Padua

Concettina Guerra is joint faculty at georgia Tech and Padua, and we started this collaboration through her. An assistant professor at the University of Padua, Matteo Comin, has implemented a research idea that originated from other research carried out in the projects. We are still actively pursuing this collaboration.

Other Collaborators or Contacts

Hailin Jin, Adobe Research
Pablo Fernandez, University of Alcala, Madrid, Spain
Prof. Bruce Walker, Georgia Tech
Prof. Greg Turk, Georgia Tech
Prof. Tucker Balch, Georgia Tech
Prof. James M. Rehg, Georgia Tech
Prof. Eric Vigoda, Georgia Tech
Prof. Teck Khim Ng, DSTA Singapore & University of Singapore
Prof. Wolfram Burgard, of the University of Freiburg
Emanuele Menegatti, University of Padua
Concettina Guerra, University of Padua and Georgia Tech
Matteo Comin, University of Padua

Activities and Findings

Research and Education Activities:

In 2010, the final year of the project, we have expanded on the large-scale graphical model inference work, which constituted the dissertation work of Kai Ni, and also continued the collaboration with Pablo Fernandez on the visibility-based correspondence.

Research activities:

In the last year of this grant, the remaining funds were used to fund Kai Ni, who completed his final semester at Georgia Tech in the Fall of 2010. Kai developed and implemented a completely hierarchical version of a divide-and-conquer to simultaneous localization and mapping (SLAM) in robotics. He showed that the original SLAM graph can be recursively partitioned into multiple-level submaps using the nested dissection algorithm, which leads to the cluster tree, a powerful graphical model representation that has not been used in this context before. By employing the nested dissection algorithm, the algorithm eliminates dependencies between two subtrees, and the optimization of the original SLAM graph can be done using a bottom-up inference along the corresponding cluster tree. To speed up the computation, we also introduced a base node for each submap and use it to represent the rigid transformation of the submap in the global coordinate frame.

As the final keystone of his dissertation work, Kai also applied this hierarchical nested

dissection scheme within the context of Structure
from Motion, the problem of inferring 3D scene structure from a unordered collection of images. This domain, while superficially similar to the SLAM problem at a mathematical level, comes with a very different set of assumptions and lacks any odometric information between cameras. In addition, because cameras are projective devices, one can easily have degeneracy problems when naively sub-dividing a large problem as in the SLAM case. Hence, one of the crucial problems that needed to be solved in this context is the how to generate good partitionings for SfM that divide the problem into fully constrained sub-problems. Kai Ni and the PI's proposed solution is 'HyperSfM',
a principled way to recursively divide an SfM problem using a hypergraph representation, in which finding edge separators yields the desired ?nested-dissection? style tree of nonlinear sub-problems.

Collaborative Research:
Kai Ni and the PI also continued the very fruitful collaboration with Pablo Fernandez Alcantarilla of Alacala University in Madrid, on the visibility-based correspondence analysis. As discussed before, a crucial step in many vision based applications, such as localization and structure from motion, is the data association between a large map of known 3D points and 2D features perceived by a new camera. In this collaboration, we investigated a novel approach to predict the visibility of known 3D points with respect to a query camera in large-scale environments. In our approach, we model the visibility of each 3D point with respect to a camera pose using a memory-based learning algorithm, in which a distance metric between cameras is learned in an entirely non-parametric way.

Educational Activities:
Kai Ni has now finished and successfully defended his dissertation, and started work for Microsoft in a group that is developing a product in the area of large-scale reconstruction.

Outreach Activities:
The GTSAM toolbox, discussed in more detail below, was released to the research community at IROS 2010, free for government and research use. Commercial applications still have to obtain a license from the Georgia Tech licensing office. Since then (November 2010) it has been downloaded over 50 times by researchers from all over the world. Future releases of GTSAM will include the TSAM work developed under the current award, but some work is still needed to make it more accessible to non-specialists.

In 2009 we have continued the large-scale graphical model work, visibility-based correspondence, and developed a new correspondence method in collaboration with Adobe research.

Research activities:

Christian Potthast was funded in part from the award in order to investigate extensions of large-scale graphical SLAM to multi-robot configurations. He spent the spring and summer on implementing a distributed implementation of the 2005 AAAI work, which still had a centralized processing paradigm. The work executed under the CAREER award led to new ideas which we are now continuing under a different (ARL-funded) grant. Basically, the idea is to remove a central point of failure to obtain a truly distributed SLAM algorithm. We are also investigating the correspondence problem in this application and are looking towards two-point RANSAC schemes.

Christian also worked on the general GTSAM code-based, where GTSAM stands for ?Georgia Tech Smoothing and Mapping?, a completely factor-graph-based software toolbox that used in both research and teaching (see below).

Collaborative Research:
The visibility-based correspondence work (see below, in 2008) has been submitted to

ICRA in a robotics context, and accepted, and submitted to ECCV in a computer vision context. This work was executed by Kai Ni in collaboration with Pablo Fernandez of Alacala University in Madrid, Spain.

Kai Ni has also written up an improved version of the GROUPSAC algorithm, a new RANSAC variant that was developed in collaboration with Adobe research. The basic idea there is to use perceptual grouping of extracted features and a new probabilistic model in order to speed up RANSAC.

Educational Activities:
The PI continues to integrate his research in graphical models for SLAM and Structure from Motion with his teaching activities. A substantially revised version of the 3D reconstruction course, developed in 2008, is being taught now in the 2010 Spring Semester. In contrast to the first instantiation of the course, the second time around the students actually got to install and use the GTSAM toolbox as part of an assignment, and several project teams in the class are using GTSAM as the basis for their project.

In 2008:

- Career Development

The PI?s career goals are still very much in line with the original research goals stated in the CAREER grant proposal, particularly with regards to examining solutions to large-scale robot mapping (SLAM) and 3D reconstruction (SFM) problems. As can be seen below, both the educational and research project line up well with those stated goals. Especially the work on graphical models and SLAM has enjoyed widespread visibility in the community and has impacted thinking in this area. To further awareness of the computational structure that large-scale SLAM and SFM problems share, the PI?s plans include both educational and outreach, respectively through a new graduate course and a book on the subject.

- Course development: CS 8803, 3D RECONSTRUCTION AND MAPPING IN COMPUTER VISION, ROBOTICS, AND AUGMENTED REALITY

The PI developed a graduate course on the state of the art in 3D reconstruction, with an emphasis on graphical models and tying in concepts from machine learning. This course was offered for the first time in Spring 2009, when about 25 graduate students took it, and will be taught two years in a row, once every two years after that. The intent is that it will form the basis for a book Prof. Dellaert is writing on the subject.

- Visibility-based Correspondence

Activities: In 2008, the PI and his graduate student Sang Min Oh started developing a new idea for speeding up and improving feature correspondence in the context of localization in environments for which pre-built 3D maps are available. The 3D models are in the form of a point cloud, and the basic idea is to learn and associate the pattern of visibility for each 3D point, to be used at run-time. This idea was developed to improve the performance in another NSF-funded project for localization for the visually impaired, with co-PI Bruce walker. It was worked out in collaboration with a post-doc in that project, Gian-Luca Mariottini, and a visiting scientist from Spain, Pablo Fernandez, from the University of Alcala.

- Photometric Correspondence in 4D Cities

Activities: A second new thread we explored in 2008 is to look at pixel correspondences over time in a 4D reconstruction of an urban environment. Based on the 4D models developed in the 4D Cities project, also NSF-funded with the same PI, we established

pixel-by-pixel correspondence between images of the same urban structures/buildings but taken at different times, possibly decades apart. The goal was then to infer both the photometric properties of the buildings as well as the color transfer characteristics of the cameras with which the images were taken. Detailed correspondence was obtained by refining the rough 3D correspondence based on the photometric models.

- Large-scale Graphical SLAM

Activities: in 2008 the PI worked with visiting scientist Christian Potthast to create a C++ library that implements the ideas developed above, specifically factor-graph-based problem definition and optimization. The goal is to also make the library available in MATLAB via a foreign-language interface to promote wide dissemination to the research and robotics community.

In 2007:

- Large-scale Graphical SLAM

Background: in 2006, the PI started a novel thread of research which is not about correspondence per se, but about efficient optimization when correspondence is known, prompted by the computational complexity of large-scale structure from motion problems encountered earlier in this project, This work was started with two unfunded M.Sc. students, Peter Krauthausen and Alex Kipp, and continued on with Michael Kaess, who has been funded on the current project last year as well, and Ananth Ranganathan. Michael and Ananth extended the 2005 work which appeared in AAAI and RSS, towards highly efficient incremental structure from motion and/or robot mapping (both problems are essentially the same optimization problem).

Activities: In 2007, PhD students Michael Kaess (funded from other sources) and Kai Ni (funded in Fall 2007 from the grant) continued this very fruitful research direction. Michael attacked the correspondence problem in incremental SAM (iSAM). Kai was funded to extend the work towards very large-scale environments, using a divide and conquer approach.

- Sampling-based Topological Mapping

Activities: In 2007, Ph.D. student Ananth Ranganathan wrapped up the research on topological mapping for robots, based on Markov chain Monte Carlo (MCMC) techniques. Background: earlier, we have shown that the space of topologies is equivalent to the intractably large space of set partitions on the set of available measurements. The combinatorial nature of the problem, however, is overcome by computing an approximate, sample-based representation of the posterior using MCMC.

- Group theory and correspondence

Background: The PI investigated the use of group theory in correspondence-less structure from motion, very much in line with the proposed research theme. Ph.D. students Josh Jones and Ananth Ranganathan re-implemented an algorithm from Makadia and Daniilidis that uses spherical harmonics in order to impose the fundamental constraint between two images in a correspondence-free way. This was a speculative research thread in line with the CAREER proposal, primarily funded by a seed grant by DSTA.

Activities: In 2007 Ananth Ranganathan and PI wrapped up this research thread by a final report to DSTA. We have also started a collaboration with the university of Padua on investigating the use of spherical harmonics to bio-informatics problems.

January 2007

- Sampling-based Topological Mapping

Activities: The PI continued the successful work with Ph.D. student Ananth Ranganathan on
a novel approach to topological mapping for robots, based on Markov chain Monte Carlo (MCMC) techniques. Earlier, they have shown that the space of topologies is equivalent to the intractably large space of set partitions on the set of available measurements. The combinatorial nature of the problem, however, is overcome by computing an approximate, sample-based representation of the posterior using MCMC. In 2006, we have concentrated on an efficient filtering implementation, as well as incorporating novel probabilistic object recognition techniques to recognize locations in an indoor environment. The latter work is still in its infancy and will be reported on next year.

- MCMC-based association of labels with tracked segments

Activities: In a collaboration with Prof. James M. Rehg and jointly advised student, Sangmin Oh, the PI has further investigated MCMC as a tool for inference in behavior recognition. Tracks of honeybees, obtained through trackers developed in the now finished NSF Biotracking project, were segmented into distinct insect behaviors by modeling the behaviors through switching linear dynamic system (SLDS) models. Because inference is intractable in these SLDS models, MCMC was evaluated as an approximate method to accomplish the association between labels and segments.

- Large-scale Graphical SLAM

Activities: prompted by the computational complexity of large-scale structure from motion problems encountered earlier in this project, the PI has started a novel thread of research which is not about correspondence per se, but about efficient optimization when correspondence is known. This work was started with two unfunded M.Sc. students, Peter Krauthausen and Alex Kipp, and continued on with Michael Kaess, who has been funded on the current project last year as well, and Ananth Ranganathan. Michael and Ananth extended the 2005 work which appeared in AAAI and RSS, towards highly efficient incremental structure from motion and/or robot mapping (both problems are essentially the same optimization problem).

- Correspondence between 3D and 2D

Activities: With a new incoming student, John Rogers, the PI investigated the state in the art of associating 3D features in laser-rangefinder-based models with imagery taken from the same scene. While the eventual goal is to use MCMC-based inference in this context, this was an initial exploration where the scope was limited to implementing one of the existing methods.

- Guided Matching

Activities: After working with a M.Sc. student for a semester in the 4D Cities project, the PI decided to fund the student (Hunter McEwan) from the current grant in order to investigate
guided matching to establish the correspondence between two or more images that have already been placed in a 3D scene. This was an early exploration and the student was relatively inexperienced, and this particular thread did not yield any significant advances.

- Group theory and correspondence

Activities: The PI further investigated the use of group theory in correspondence-less

structure from motion, very much in line with the proposed research theme. Ph.D. students Josh Jones and Ananth Ranganathan have re-implemented an algorithm from Makadia and Daniilidis (at UPenn) that uses spherical harmonics in order to impose the fundamental constraint between two images in a correspondence-free way. This is a highly speculative research thread in line with the CAREER proposal, but is primarily funded by a seed grant by DSTA.

February 2006

- Vision-based Localization and Mapping for Mobile Robots

With my student Ananth Ranganathan I have investigated a novel approach to robotic mapping. Mapping an unknown and un-instrumented environment is one of the foremost problems in robotics. For this purpose, both metric and topological maps have been explored in depth as viable representations of the environment. Rather than attempting to build very detailed, metrically correct maps, we propose to build larger-scale topological maps. However, as stated in the proposal, the combinatorial nature of topological maps so far precluded a principled Bayesian inference approach, in contrast to metric mapping in which probabilistic methods represents the state of the art. We have established a theoretical framework in which topological maps are isomorphic to set partitions over the space of landmarks, and developed an MCMC sampler in the space of set partitions, along with algorithms to incorporate odometry in an efficient way. In particular, we developed the concept of Probabilistic Topological Maps (PTMs), a sample-based representation that approximates the posterior distribution over topologies given available sensor measurements. We have shown that the space of topologies is equivalent to the intractably large space of set partitions on the set of available measurements. The combinatorial nature of the problem, however, is overcome by computing an approximate, sample-based representation of the posterior. The PTM is obtained by performing Bayesian inference over the space of all possible topologies and provides a systematic solution to the problem of perceptual aliasing in the domain of topological mapping.

In a second development, and as proposed, my student Michael Kaess and I have attacked the loop closing problem in robotics. In the context of visual simultaneous localization and
mapping (or SLAM) the 'landmarks' in the environment are really 3D points in space, and they are measured by a camera or a set of cameras ? as in our work. The problem of loop closing is then the problem of realizing that the robot came back to the same place, and that the measurements taken by the camera correspond to the same 3D features seen earlier. Michael and I have framed this problem as establishing a set partitions over 3D measurement tracks. If tracks taken at different times are grouped together, they are assumed to be measuring the same 3D feature. The probability of this event can be systematically derived and efficiently computed by integrating out both the robot trajectory and the 3D feature position, in a process called Rao-Blackwellization.

- Biotracking and Sequential Monte Carlo Techniques

As proposed, we have extended the success of our MCMC approach to multi-target tracking to applications in which correspondence problems play a large role. This was done in collaboration with Zia Khan, a former student that was co-advised by Tucker Balch
and myself. In many multitarget-tracking applications in computer vision, a detection algorithm provides locations of potential targets. Subsequently, the measurements are associated with previously estimated target trajectories in a data association step. The output of the detector is often imperfect and the detection data may include multiple, split measurements from a single target or a single merged measurement from several targets. To address this problem, we introduced a multiple hypothesis tracker for interacting targets that generate split and merged measurements. The tracker is based on an efficient

Markov chain Monte Carlo based auxiliary variable particle filter. The particle filter is Rao-Blackwellized such that the continuous target state parameters are estimated analytically, and an MCMC sampler generates samples from the large discrete space of data associations.

In a second development, my student Grant Schindler and I have investigated the use of a 'configuration of parts' model to tracking non-rigid objects in a video stream. Specifically, we were interested in tracking non-rigid insects. Like other tracking targets, such as people, insects are physically composed of multiple parts that flex and bend with respect to each other. We wanted to model this flexible motion, which is hypothesized to improve the performance of our tracker and increase the richness of the data that can be used for subsequent analysis. As such, we adopted a model that incorporates an object's individual parts, modeling the joint configuration of the parts as a whole, and modeling the appearance of each part individually. We have investigated how to efficiently incorporate such a model into a particle filter by treating the shape analytically and sampling only over pose, a process commonly known as Rao-Blackwellization, and used in several of our other approaches described above.

- Group theory and correspondence

In a novel thread, spawned by a synergistic seed grant from DSTA, Singapore, I have been looking into how group theory can provide a deeper insight into the problem of correspondence. A central theme in my CAREER proposal was the use of MCMC to sample over the space of correspondences. Group theory, the mathematical theory of transformations of objects, provides a structure to this space. In particular, the symmetric group of permutations turns out to have a rich internal structure that can perhaps be exploited to build efficient samplers.

I have worked on this with two students: Shakti Kamal and Josh Jones. With Shakti I have primarily looked at the work by Dan Rockmore, on generalized Fourier transforms in non-Abelian groups. With Josh I have looked at the work by Makadia, Geyer, and Daniilidis, who used harmonica analysis on the group of 3D rotations to establish correspondence between two images.

Findings:

In 2010:

We submitted and published the results of Kai Ni's dissertation work as applied in the SLAM domain to IROS 2010 under the name TSAM 2.0 (Tectonic SAM 2.0). In this paper, Kai's showed that TSAM 2, by virtue of introducing base nodes and optimizing them, rather than the actual submap variables in every iteration, significant computational savings can be achieved. He demonstrated experimentally that the algorithm is not only exact but also much faster than alternative approaches, in both simulations and real-world experiments.

In addition , the HyperSfm work has been submitted to ICCV 2011 and is currently under review. in the paper, Kai shows that after partitioning, a bottom-up computation pass solves the SfM problem robustly (by having fully constrained sub-problems) and efficiently (because most nonlinear error is removed at lower levels of the tree). The performance of the algorithm is demonstrated for various indoor and outdoor standard data-sets.

The collaboration with Pablo Fernandez Alcantarilla resulted in another ICRA paper, presented in May 2011 in Shanghai, in which it was shown that by fully exploiting the geometric relationships between the 3D map and the camera poses, as well as the related appearance information, the resulting visibility prediction is much more robust and efficient than conventional approaches. We demonstrated the performance of the new

algorithm on a large urban 3D model in terms of both speed and accuracy.

In 2009:

The visibility-based correspondence idea works very well for large-scale structure from motion applications in addition to visual SLAM/localization, for which it was first developed.

The GROUPSAC algorithm is a much faster approach than the vanilla RANSAC approach. The paper showing experimental results was submitted and accepted to CVPR 2009.

In 2008:

- Visibility-based Correspondence

Findings: Using memory-based learning techniques, the idea of visibility-based correspondence allows for a quick pruning of putative correspondences at run-time based on a prior pose estimate, drastically cutting down on the time needed to throw out outlier correspondences. Results are very promising, and a publication draft is currently under review.

- Photometric Correspondence in 4D Cities

Findings: This work was done by graduate student Mingxuan Sun, and was followed up with a re-coloring project co-supervised by Prof. Greg Turk. The pixel-by-pixel correspondence and the resulting 4D models with changing texture maps produced very nice visualizations of large-scale 4D structures over time. However, the student in question did not produce results of sufficient quality to convince reviewers of the merit of the idea. The only publication resulting from this was a SIGGRAPH sketch. While the idea has merit, much more work is needed, with more complex models, to develop it into something worthwhile. The PI has decided to not pursue this idea further at this time, also because the student carrying out the work is no longer with his research group.

- Large-scale Graphical SLAM

Findings: the C++ library ?gtsam?, (Georgia Tech Smoothing And Mapping) is currently under development and is being beta-tested in a course developed and given by the PI. We expect to release the C++ library and accompanying MATLAB toolbox in the course of the 2009 calendar year.

In 2007:

- Large-scale Graphical SLAM

Findings: Michael developed computationally efficient ways to compute covariance matrices needed to perform data-association in an incremental SLAM setting. This work was published at ICRA 2007 in Rome. Furthermore, using probabilistic models first developed in the PI?s doctoral thesis, Michael attacked the loop-closing problem in Smoothing and Mapping. Both a conference paper and a journal paper are currently under review.

Kai?s work resulted in two papers in first-tier conferences, one in ICRA and the other in ICCV, where the work was very well received. In these papers he develops the idea of breaking up the large graph of constraints inherent to the SLAM problem, and attacking each subgraph in turn. This leads to an efficient, out-of-core method that does not have to keep the entire problem in memory at all times. The different subgraphs are then combined in a final step to yield the global solution. While the method was developed to

attack very-large scale problems, a nice bonus result was that the resulting method is in fact faster than previous methods, as well.

- Sampling-based Topological Mapping

Findings: Ananth has now incorporated novel probabilistic object recognition techniques to recognize locations in an indoor environment. This work was presented at the RSS 2007 conference in Atlanta. In addition, he has written up his dissertation and will defend his PhD thesis in February 2008.

- Group theory and correspondence

Findings: while group theory is still believed to be a interesting avenue of exploration, we concluded that the Makadia work is really not practical in its current form, and not competitive with other state of the art methods such as RANSAC. The bio-informatics thread, however, seems quite promising, and we will be continuing this work in the future.

January 2007

- Sampling-based Topological Mapping

Findings: An efficient filtering-based technique was developed for constructing topological
maps in an incremental but robust manner, and was presented at the ICRA 2006 conference. The algorithm uses laser range data and odometry in a Rao-Blackwellized Particle Filter (RBPF) setting to perform inference in the joint space of topologies and landmark locations. Experiments were presented on data gathered from robot runs to validate the algorithm.

- MCMC-based association of labels with tracked segments

Findings: A novel inference method for SLDS was developed and presented at the prestigious AAAI conference. The proposed method provides a concrete framework to a variety of research communities. First, it can characterize the accuracy of deterministic approximation algorithms. Secondly, it serves as an inference method that can discover true posteriors. This leads to a robust inference and learning. Finally, DD-MCMC delivers correct MAP solutions to a wide range of SLDS applications. The proposed method efficiently converges to the posterior, overcoming a potential limitation of MCMC methods.
The efficiency is achieved via the sampling space reduction by Rao-Blackwellization and the data-driven approach where temporal cues and proposal priors are introduced to construct an efficient proposal. In terms of characteristics, DD-MCMC inference method embraces data characteristics and model-based approaches. The characteristics of data are effectively learned using the temporal cues where the form of the cues are obtained from prior knowledge.

- Large-scale Graphical SLAM

Findings: At IJCAI 2007, Ananth Ranganathan presented the resulting algorithm, incremental smoothing and mapping or iSAM, at the prestigious IJCAI ?07 conference in Hyderabad, India. The algorithm achieves efficiency by updating the square-root information matrix, a factored version of the naturally sparse smoothing information matrix. One can efficiently recover the exact trajectory and map at any given time by back-
substitution. Simulation-based results for the linear case were presented, showing constant time updates for exploration tasks. The behavior of the algorithm was further evaluated in the presence of loops, as well as how the approach extends to the non-linear

case. Finally, the overall non-linear algorithm was evaluated on the standard Victoria Park data set.

- Group theory and correspondence

Findings: The group theory thread has lead to a better the understanding of permutation groups for inference on the space of explicit correspondences, and the group (really, homogenous space) of spherical harmonics for a more implicit (correspondence-free) way of doing the same inference.

February 2006

- Vision-based Localization and Mapping for Mobile Robots

One of the main findings in the MCMC-driven topological mapping is the considerable mileage to be gotten from a data-driven MCMC sampler, which considerably speeds up inference. In addition, we have found that the use of appearance also heavily constrains the set of feasible topological maps over which inference needs to be done. Specifically, we considered appearance measurements in the form of panoramic images obtained from a camera rig mounted on a robot. At the first conference on Robotics, Systems and Science, in Boston 2005, we have presented experimental results that illustrate the robustness and wide applicability of our algorithm, and generate good maps when using odometry and appearance, derived from panoramic images, as sensor measurements.

As for the loop closing work, there we have applied the algorithm to corner features extracted from laser range data. The corners are obtained by a modified version of the incremental line-fitting algorithm by extracting intersections between neighboring line segments. The findings here were less positive than expected. As with every optimization technique, EM can easily get stuck in local minima. We found that this can ? to some extent ? be avoided by applying an annealing scheme as is commonly used in the context of EM. In an ICRA 2005 ICRA conference paper we have presented the results of experiments were we have applied our algorithm to two sets of laser data. The first was recorded in our research facility, while the second data set is part of the Intel Oregon sequence that was obtained from the Robotics Data Set Repository. After applying our algorithm, we obtained the corrected maps in both cases. However, more work is needed to cope with the problems of combining EM and sampling in this context.

- Biotracking and Sequential Monte Carlo Techniques

Applying MCMC to the correspondence problem in a multi-target tracking application works well in practice. We obtained experimental results in a scenario where we track several interacting targets that generate these split and merged measurements. Because the application of visual tracking technologies to monitoring the movement of animals has important implications in the study of behavior, we applied the algorithm to tracking interacting ants in a behavioral experiment. The application presents a substantial challenge as targets deform and frequently interact, and hence merged and split measurements are common during tracking.

As far as the parts-based RBPF, we have shown that it reduced tracker failures by a factor of 2 in the experiments we have conducted. We used somewhat na?ve appearance and motion models, in part, so that we could isolate and observe more clearly the specific advantages of a parts-based model for difficult tracking tasks. Only with the addition of more sophisticated appearance models (e.g. subspace models) and motion models (e.g. switching linear dynamic systems) would we expect the tracker to perform perfectly. In summary, what we have found is that, (1) parts-constellation models can be beneficial for some tracking tasks, and (2) Rao-Blackwellization enables the efficient use of parts-based models for particle filtering.

Training and Development:

April 2011

Kai Ni successfully defended his dissertation on April 21st, and started work for Microsoft in a group that is developing a product in the area of large-scale reconstruction.

May 2009

Sang Min Oh, who has been funded under this grant, has graduated and now works for Kitware. Christian Potthast, who was a visitor in the lab but functioned as a research assistant, has successfully applied to several US grduate programs and has started in Fall 2009 at USC, under the direction of Gaurav Sukathme.

February 2009

Both Michael Kaess and Ananth Ranganathan have graduated. Michael Kaess is now a postdoc at MIT, Ananth Ranganathan is working on humanoid robots at Honda research Labs.

February 2008

Both Michael Kaess and Ananth Ranganathan are expected to graduate in Spring 2008.

January 2007

Several graduate students have published peer-reviewed papers in top-tier conferences and journals, and have further deepened their knowledge of their chosen research topic.

Ananth Ranganathan has orally presented two papers at the prestigious IJCAI ?07 conference in India.

February 2006

As discussed in the project proposal, Professor Tucker Balch have designed an undergraduate course on computational robotics and perception, that puts more emphasis on mathematics and probability than traditional robotics intro courses. We have made this into a 300-level class (juniors and higher) so that interested students can take a senior-level course either in robotics and vision. In this course, the Bayesian sampling paradigm at the core of my research activities also serves as an excellent educational vehicle, as it is a way to avoid the traditionally formula-heavy approach to statistics and probability. We are now in our second semester of teaching the course.

Several graduate students have published peer-reviewed papers in top-tier conferences and journals, and have further deepened their knowledge of their chosen research topic.

Michael Kaess has spent one semester in Freiburg, collaborating with Prof. Wolfram Burgard, and one semester as an intern at Microsoft Research, and hence has extended his research horizon considerable.

Grant Schindler has presented a paper at a workshop in Beijing, and has benefited both from attending one of the major vision conferences (ICCV 2005) as well as visits to Tsinghua University and Microsoft research, Asia.

Outreach Activities:

The PI gave a keynote talk at a graphical model workshop at IROS 2010, and as part of his participation in the workshop announced the release of the GTSAM graphical model inference toolbox to the research community. This toolbox is free for government and research use, although commercial ventures have to obtain a license from the Georgia Tech licensing office.

The PI has also given a talk on large-scale inference at the RSS area chair meeting, as well as at various universities throughout 2010 (including ETH Zurich).

The PI has given an invited talk at USC and at HRL, in which he advocated the use of graphical models in order to model and solve large-scale SLAM and SFM problems.

February 2009

The PI has participated in the 2009 SLAM Summer School at the ACFR in Sydney, Australia, to teach and disseminate the principles behind large-scale graphical SLAM. In Spring 2009 he is teaching a newly developed course on 3D reconstruction and mapping and is using this class to help develop an open source library for graphical SLAM.

February 2008

The PI has given numerous invited talks in 2007 that highlighted the work supported by the project, including at Google, Microsoft, ETH Zurich, LAAS in Toulouse, and at the University of Leuven. In addition, the PI has developed a tutorial on Monte Carlo methods in Robotics and Vision, given at the University of Liege in Belgium to a diverse audience. The tutorial is also available on-line.

January 2007

The PI gave an invited talk at the University of Maryland and at the University of California, San Diego which highlighted research supported by this award.

February 2006

The PI gave an invited talk at the department of CS at the University of Toronto on the tracking work, and invited lectures at Microsoft research and the University of Washington on MCMC. In addition, he has co-organized an ICCV tutorial on the use of MCMC in computer vision, in collaboration with Song-Chun Zhu and Zuowhen Tu of UCLA.

Michael Kaess has given talks during his time in Freiburg and as an intern at Microsoft Research.

Journal Publications

Ananth Ranganathan E. Menegatti and Frank Dellaert, "Bayesian Inference in the Space of Topological Maps", IEEE Transactions on Robotics, p. 92, vol. 22, (2006). Published,

Zia Khan, Tucker Balch, and Frank Dellaert, "MCMC-Based Particle Filtering for Tracking a Variable Number of Interacting Targets", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, p. 1805, vol. 27, (2005). Published,

Z. Khan, T. Balch, and F. Dellaert, "MCMC Data Association and Sparse Factorization Updating for Real Time Multitarget Tracking with Merged and Multiple Measurements", IEEE Transactions on Pattern Analysis and Machine Intelligence, p. 1960, vol. 28, (2006). Published,

Sang Min Oh and James M. Rehg and Tucker Balch and Frank Dellaert, "Learning and Inferring Motion Patterns using Parametric Segmental Switching Linear Dynamic
Systems", International Journal of Computer Vision, p. , vol. , ( ). Accepted,

Frank Dellaert and Michael Kaess, "Square Root SAM: Simultaneous Location and Mapping via Square Root Information Smoothing", International Journal of Robotics Research, p. , vol. , (2006). Accepted,

Kaess, M; Ranganathan, A; Dellaert, F, "iSAM: Fast incremental smoothing and mapping with efficient data association", PROCEEDINGS OF THE 2007 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1-10, p. 1670, vol. , (2007). Published,

Michael Kaess, Ananth Ranganathan, and Frank Dellaert, "iSAM: Incremental Smoothing and Mapping", IEEE Transactions on Robotics, p. 1365, vol. 24, (2008). Published,

Ananth Ranganathan and Frank Dellaert, "Online probabilistic topological mapping", International Journal of Robotics Research, p. 755, vol. 30, (2011). Published,

Michael Kaess, Frank Dellaert, "Probabilistic structure matching for visual SLAM with a multi-camera rig", Computer Vision and Image Understanding, p. 286, vol. 114, (2010). Published,

Books or Other One-time Publications

Kai Ni, Hailin Jin, and Frank Dellaert, "GroupSAC: Efficient Consensus in the Presence of Groupings", (2009). Conference Paper, Published
Collection: IEEE International Conference on Computer Vision (ICCV)
Bibliography: GroupSAC: Efficient Consensus in the Presence of Groupings, Kai Ni, Hailin Jin, and Frank Dellaert, IEEE International Conference on Computer Vision (ICCV), 2009

Pablo Fern?ndez Alcantarilla, Sang Min Oh, Gian Luca Mariottini, Luis Miguel Bergasa, and Frank Dellaert, "Learning Visibility of Landmarks for Vision-Based Localization", (2010). Conference Paper, Accepted
Collection: IEEE International Conference on Robotics and Automation (ICRA)
Bibliography: Learning Visibility of Landmarks for Vision-Based Localization, Pablo Fern?ndez Alcantarilla, Sang Min Oh, Gian Luca Mariottini, Luis Miguel Bergasa, and Frank Dellaert, IE

Web/Internet Site

URL(s):
www.cc.gatech.edu/4d-cities/dhtml/index.html
Description:

Other Specific Products

Product Type:

Conference Paper

Product Description:

Visibility Learning for Large-Scale Urban Environment, Pablo Fern?ndez Alcantarilla, Kai Ni, Luis Miguel Bergasa, and Frank Dellaert, IEEE International Conference on Robotics and Automation (ICRA), 2011

Sharing Information:

Presented at ICRA 2011

Product Type:

Conference Paper

Product Description:

Multi-Level Submap Based SLAM Using Nested Dissection, Kai Ni and Frank Dellaert, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010

Sharing Information:

Presented at IROS 2010

Contributions

Contributions within Discipline:

In 2010, we prepared another ICRA paper on visibility, and had an IROS paper on TSAM2

In 2009, we had one ICCV paper on GROUPSAC, and one ICRA paper on visibility-based correspondence.

February 2008:

In addition to the journals listed in the previous section, the PI and his collaborators have published the following peer-reviewed conference papers related to the award topic:

? Semantic Modeling of Places using Objects, Ananth Ranganathan and Frank Dellaert, Robotics: Science and Systems, 2007
? Out-of-Core Bundle Adjustment for Large-Scale 3D Reconstruction, Kai Ni, Drew Steedly, and Frank Dellaert, IEEE International Conference on Computer Vision (ICCV), 2007
? Fast 3D Pose Estimation With Out-of-Sequence Measurements, Ananth Ranganathan, Michael Kaess, and Frank Dellaert, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2007
? iSAM: Fast Incremental Smoothing and Mapping with Efficient Data Association, Michael Kaess, Ananth Ranganathan, and Frank Dellaert, IEEE International Conference on Robotics and Automation (ICRA), 2007
? Tectonic SAM: Exact; Out-of-Core; Submap-Based SLAM, Kai Ni, Drew Steedly, and Frank Dellaert, IEEE International Conference on Robotics and Automation (ICRA), 2007

? Large-scale Graphical SLAM

Contributions: The iSAM algorithm is a novel approach to the problem of simultaneous localization and mapping (SLAM) based on incremental smoothing, that is suitable for real-time applications in large-scale environments. The main advantages over filter-based algorithms are that it solves the full SLAM problem without the need for any approximations, and that it does not suffer from linearization errors. Furthermore, our approach allows access to the exact covariances, as it does not suffer from under-estimation of uncertainties, which is another problem inherent to filters.

? Sampling-based Topological Mapping

Contributions: Efficiently obtaining topological maps not only solve the problem of perceptual aliasing but also provide the basis for the construction of global metric maps. The use of a Dirichlet process prior on the landmark labels is also a significant and novel contribution of this work. The use of a data-driven proposal distribution to overcome degeneracy in the particle filter was another contribution. Finally, in 2007 we made the first steps towards incorporating object recognition in a natural way within the same framework.

? MCMC-based association of labels with tracked segments

Contributions: To the best of their knowledge, the PI and his collaborators believe that the work on MCMC-based label-association is the first to explore Data-Driven MCMC techniques for switching linear dynamics systems. Moreover, the proposed framework is broadly-applicable to problems in time series and behavior modeling.

January 2007

? Sampling-based Topological Mapping

Contributions: The ICRA 06 paper was nominated for best conference paper. The efficiently obtained topological maps not only solve the problem of perceptual aliasing but also provide the basis for the construction of global metric maps. The use of a Dirichlet process prior on the landmark labels is also a significant and novel contribution of this work. The use of a data-driven proposal distribution to overcome degeneracy in the particle filter was another contribution.

? MCMC-based association of labels with tracked segments

Contributions: To the best of their knowledge, the PI and his collaborators believe that the work on MCMC-based label-association is the first to explore Data-Driven MCMC techniques for switching linear dynamics systems. Moreover, the proposed framework is broadly-applicable to problems in time series and behavior modeling.

? Large-scale Graphical SLAM

Contributions: The iSAM algorithm is a novel approach to the problem of simultaneous localization and mapping (SLAM) based on incremental smoothing, that is suitable for real-time applications in large-scale environments. The main advantages over filter-based
algorithms are that it solves the full SLAM problem without the need for any approximations, and that it does not suffer from linearization errors. Furthermore, our approach allows access to the exact covariances, as it does not suffer from under-estimation of uncertainties, which is another problem inherent to filters.

? Group theory and correspondence

Contributions: The PI has identified the spherical harmonics as a better way to compare and match locations in mobile robot mapping. In addition, there is some promise of using exactly the same machinery in an entirely different application, namely that of protein-ligand surface matching.

In addition to the journals listed in the previous section, the PI and his collaborators have published the following peer-reviewed conference papers related to the award topic:

? Fast Incremental Square Root Information Smoothing, Michael Kaess, Ananth

Ranganathan, and Frank Dellaert, International Joint Conference on Artificial Intelligence (IJCAI), 2007
? Loopy SAM, Ananth Ranganathan, Michael Kaess, and Frank Dellaert, International Joint Conference on Artificial Intelligence (IJCAI), 2007
? Exploiting Locality by Nested Dissection For Square Root Smoothing and Mapping, Peter Krauthausen, Frank Dellaert, and Alexander Kipp, Robotics: Science and Systems, 2006
? A Rao-Blackwellized Particle Filter for Topological Mapping, Ananth Ranganathan and Frank Dellaert, IEEE International Conference on Robotics and Automation (ICRA), 2006
? A Multifrontal QR Factorization Approach to Distributed Inference applied to Multi-robot Localization and Mapping, Frank Dellaert, Alexander Kipp, and Peter Krauthausen, National Conference on Artificial Intelligence (AAAI), 2005
? Data-Driven MCMC for Learning and Inference in Switching Linear Dynamic Systems, Sang Min Oh, James M. Rehg, Tucker Balch, and Frank Dellaert, National Conference on Artificial Intelligence (AAAI), 2005
? Learning and Inference in Parametric Switching Linear Dynamic Systems, Sang Min Oh, James M. Rehg, Tucker Balch, and Frank Dellaert, IEEE International Conference on Computer Vision (ICCV), 2005
? A Markov Chain Monte Carlo Approach to Closing the Loop in SLAM, Michael Kaess and Frank Dellaert', IEEE International Conference on Robotics and Automation (ICRA), 2005
? Multitarget Tracking with Split and Merged Measurements, Zia Khan, Tucker Balch, and Frank Dellaert, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005
? Data driven MCMC for Appearance-based Topological mapping, Ananth Ranganathan and Frank Dellaert, Robotics: Science and Systems, 2005
? A Rao-Blackwellized Parts-Constellation Tracker, Grant Schindler and Frank Dellaert, ICCV Workshop on Dynamical Vision; International Conference on Computer Vision, 2005

February 2006

A principled solution to topological mapping is a crucial component in scaling up robotic mapping to large environments. The problem of loop closing is important because accumulated measurement error poses an inherent limitation on the accuracy by which environment models can be recovered. Identifying which measurements have already been seen is essential in order to obtain accurate, globally correct models. Any incremental method that tracks features and instantiates new tracks in the standard way will end up representing the same environment features multiple times.

Multi-target tracking is a difficult and important problem that has not been satisfactorily solved under general circumstances. The algorithms we developed are heavily cited and will inspire new developments, in particular those based on MCMC, a central component in
this project.

Contributions to Other Disciplines:

2010

The PI has now published a journal paper on protein structure matching in collaboration with the University of Padua, directly inspired by the work done under this award.

In addition, the work done by Kai Ni on using nested dissection-style divide and conquer is applicable outside the domains of robotics and computer vision, wherever there is a large-scale *non-linear* optimization problem to be solved.

January 2007

The PI has identified the interesting possibility of using spherical harmonics and correspondence-free pose estimation in order to attack the scientifically important problem of protein-ligand matching. To that end, he has set up a collaboration with Concettina Guerra, a well known bio-informatics researcher that is jointly appointed between Georgia Tech and the university of Padua. It is expected that this will lead to a new research project possibly funded by NIH and/or NSF.

February 2006

In general, Markov chain Monte Carlo sampling has only been used extensively in theoretical physics, statistics and computational complexity (theory). My group is one of the few (another is that of Song Chun Zhu at UCLA) that use MCMC in large-scale vision and robotics applications, and we are establishing it as a viable alternative for other ? perhaps more heuristic- approaches in those disciplines. However, as such we provide an existence proof and an example that can transfer to other disciplines.

Contributions to Human Resource Development:

The PI book draft has evolved to include the use of the factor graph paradigm within the context of logic and constraint satisfaction, as a consequence of developing a graduate AI course based on and around factor graphs. This has delayed the book plans somewhat, but considerable opns up the scope towards graduate students in many related fields.

The PI is writing a book chapter on the use of graphical models as underlying traditional optimization methods which can be used in other engineering disciplines as a pedagogical devise.

MCMC as technique in engineering in science is new and relatively unexplored, and we need
a new breed of researchers intimately familiar with approximate Bayesian inference. Through
the Ph.D.-level projects that I have undertaken with my students I hope to create exactly that
type of student. In addition, I have in the past and will continue to work with a large number
of undergraduate students, as is a tradition at Georgia Tech.

Contributions to Resources for Research and Education:

The PI has developed a new course on large-scale graphical SLAM, a condensed version of which can be taught at conferences, summer schools etc. He has given versions of a tutorial based on the material at the SLAM summer school in Oxford in 2007, and in Sydney in 2009.

In addition, in Spring 2011, the PI developed a new graduate course in Artificial Intelligence, based partly on factor graphs, the insight in which is a direct consequence of the current award. The class has met with overwhelmingly positive responses from the students.

Contributions Beyond Science and Engineering:

Conference Proceedings

Schindler, G;Krishnamurthy, P;Lublinerman, R;Liu, YX;Dellaert, F, Detecting and matching repeated patterns for automatic geo-tagging in urban environments, "JUN 23-28, 2008", 2008 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, VOLS 1-12, : 925-931 2008

Ni, K;Steedly, D;Dellaert, F, Out-of-core bundle adjustment for large-scale 3D reconstruction, "OCT 14-21, 2007", 2007 IEEE 11TH INTERNATIONAL CONFERENCE ON COMPUTER VISION, VOLS 1-6, : 2009-2016 2007

Kaess, M;Ranganathan, A;Dellaert, F, iSAM: Fast incremental smoothing and mapping with efficient data association, "APR 10-14, 2007", PROCEEDINGS OF THE 2007 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1-10, : 1670-1677 2007

Ni, K;Steedly, D;Dellaert, F, Tectonic SAM: Exact, out-of-core, submap-based SLAM, "APR 10-14, 2007", PROCEEDINGS OF THE 2007 IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS 1-10, : 1678-1685 2007

Ranganathan, A;Kaess, M;Dellaert, F, Fast 3D pose estimation with out-of-sequence measurements, "OCT 29-NOV 02, 2007", 2007 IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS, VOLS 1-9, : 2492-2499 2007

Categories for which nothing is reported:

Contributions: To Any Beyond Science and Engineering

Georgia Tech College of Computing
School of Interactive Computing

April 26, 2010

To Whom It May Concern:

I am writing this letter to certify that the School of Interactive Computing continues to support the research efforts of Dr. Frank Dellaert and that he continues to be eligible to receive support under the CAREER program. Dr. Dellaert began his full-time tenure-track position here at Georgia Tech in September, 2001 and was promoted to Associate Professor in July, 2007.

All faculty members are supported by the School to devote two-thirds of their academic year time to research. The regular teaching load for research active faculty is two courses per academic year, or one per semester. These courses typically include a graduate level class in the area of research expertise. We also include a provision for reduced load in recognition of course development activities.

In closing, I continue to monitor the career development of Dr. Dellaert and it is proceeding quite well. I attest that the PI's career development plan is supported by and integrated into the educational and research goals of the School and the organization. And, finally, I personally commit to the support and professional development of the Dr. Dellaert.

Sincerely,

Aaron F. Bobick, Professor and Chair

801 Atlantic Drive • Atlanta, GA 30332-0280 USA • Phone 404.894.4231 • Fax 404.894.0673 • www.cc.gatech.edu